EDC DAAC

Agenda

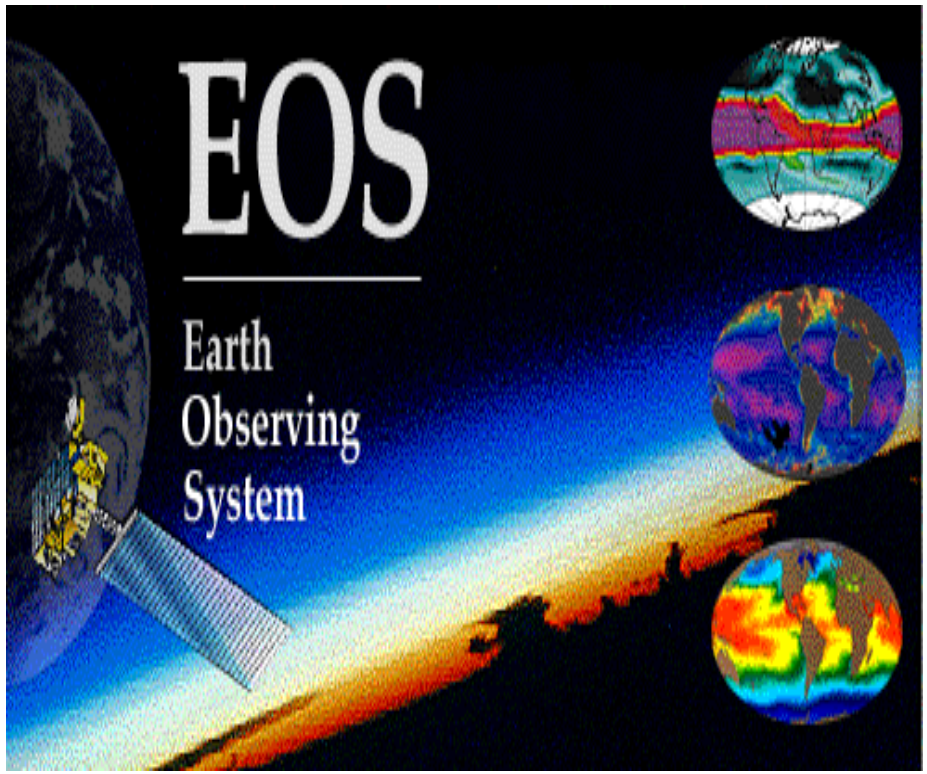
- Disclaimer
- DAAC overview/background
- Current archives
- DAAC systems
- Current ingest/distribution
- DAAC project management
- Discussion

Disclaimer

- This presentation gives an overview of the EDC DAAC, the way it works today and has worked in the past.
- *It does **NOT**, under any circumstances, represent*:
  - *Any* direction as to how the DAAC should be supported or staffed in the future.
  - *Any* recommendation, direction, or desire, technical or otherwise, as to how the EOSDIS Core Systems, or any other systems, should be administered, supported, maintained, and/or developed in the future under NASA's ESDIS Maintenance and Development (EMD) Contract.

DAAC Overview

Our vision is to combine NASA missions with USGS expertise to support the science community and make products available to, and useable by, the public, while enabling a smooth transition to the National Archive and pushing the state of the art in data center practice.



Background

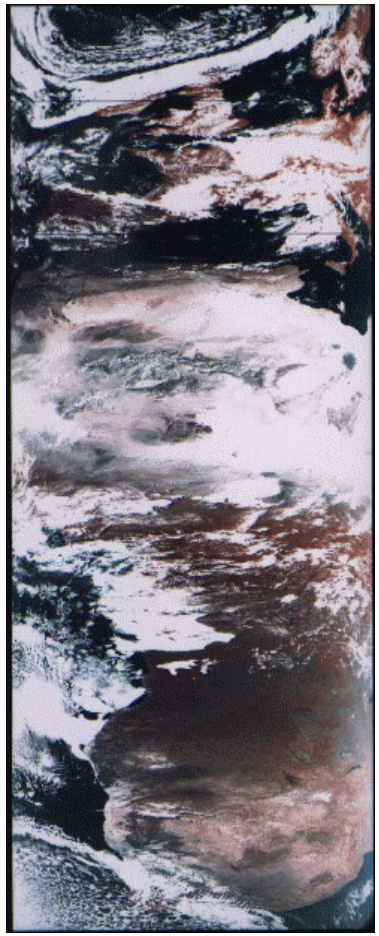
- EDC's participation with NASA's "experimental land remotely sensed data", as an "active short- and long-term archive" began with MOU in 1988.
- EOS Support Program established with NASA funding to:
  - participate in planning, definition, design of EOSDIS Core System (ECS);
  - execute the short-term archive (EDC DAAC); and
  - provide science support activities (Pathfinder, AVHRR).
- A fourth MOU responsibility, Long Term Archive (LTA), is USGS funded and not managed from the EOS Support Program.
- Program start in 1988, V0 operational in 1994, V1 (ECS) operational in 1999.

Current Archives

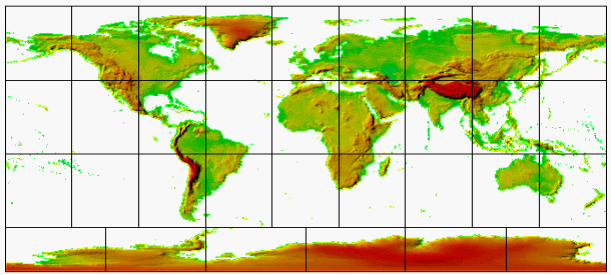
- Version 0 data
  - Implemented to:
    - Shake out processes required for Terra, Aqua, Chem, etc.
    - Provide useful data sets to users early
    - Get users used to coming to the DAACs for good data
  - Includes miscellaneous data collections available from DAAC web page or EOS Data Gateway (EDG)
    - http://edcdaac.usgs.gov/
    - http://edcimswww.cr.usgs.gov/pub/imswelcome
  - Approximately 18 TB of data in archive

Pre-EOS Data Sets

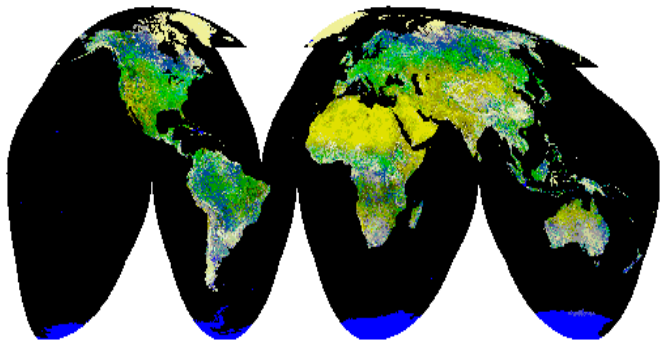
- AVHRR Orbital Stitches



- GTOPO30 Global DEM

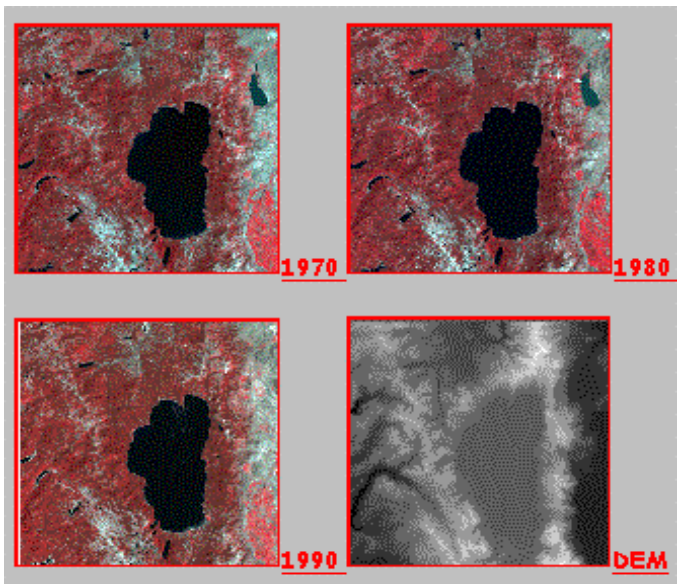


- Global AVHRR Composites

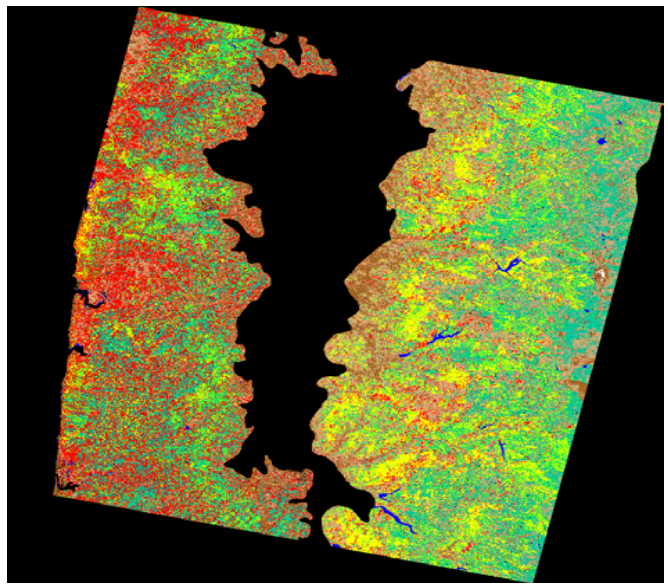


Pre-EOS Data Sets

- Landsat Pathfinder data
  - North American Land Cover (NALC)

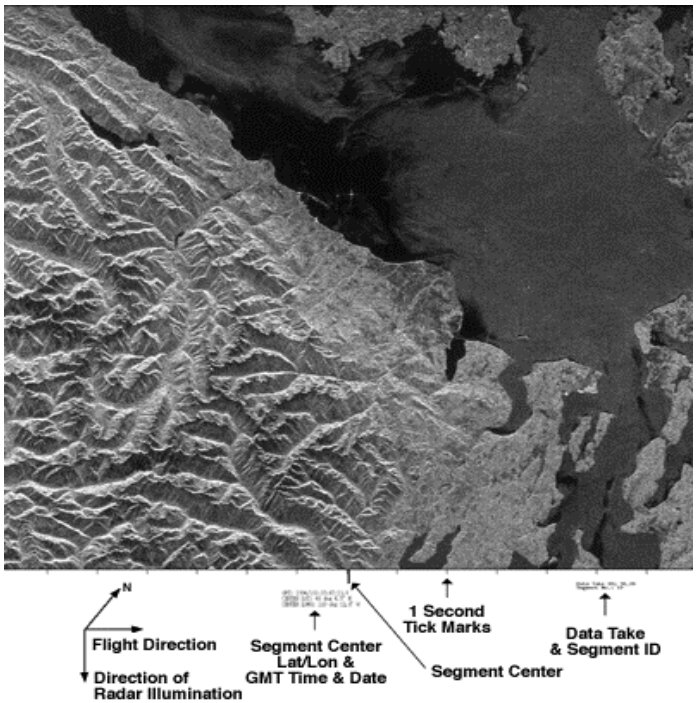


- Global Landcover Test Sites (GLCTS)

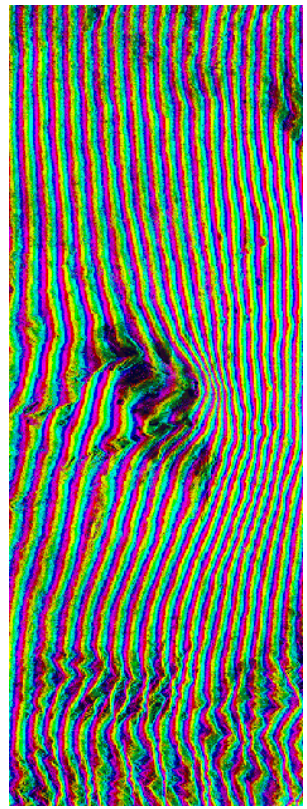


Pre-EOS Data Sets

- SIR-C Survey Data



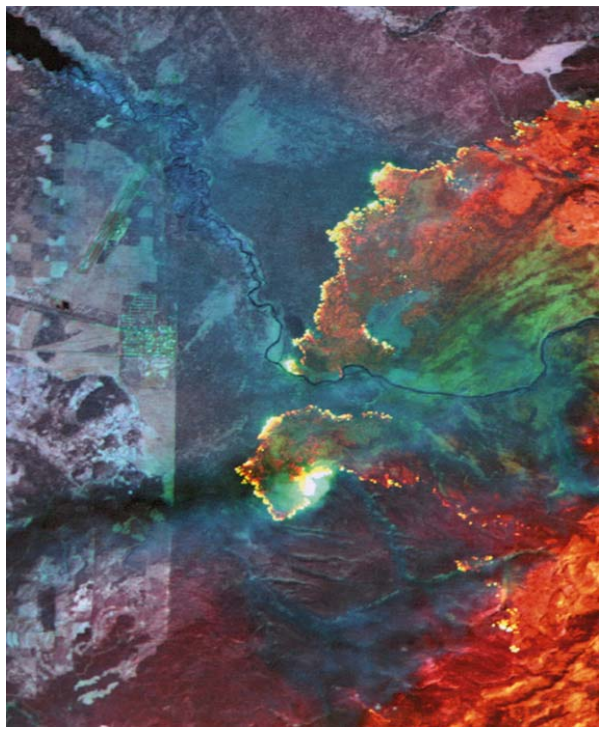
- SIR-C Precision Data



Pre-EOS Data Sets

- NASA Landsat Data Collection (NLDC)

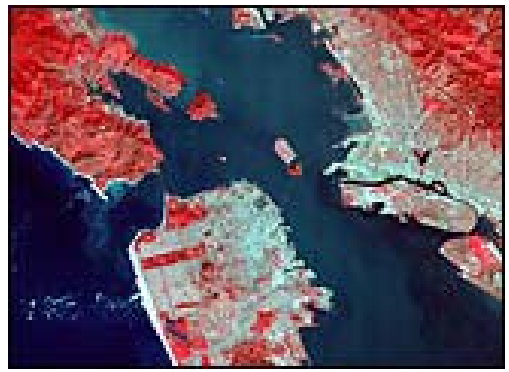
- Aircraft Scanner Data
  - TIMS/NS-001/TMS, ASAS



Current Archive

- Landsat 7
  - Primary archive and distribution for Landsat 7 L0R products
  - Currently over 86 TB of data in archive
  - Ingest up to 154 GB/day through local network
  - Distribute up to 55 GB/day of scene-based products

Current Archive



- ASTER
  - First data at EDC in March 2000.
  - Archive contains multiple data types
    - Public release began in November 2000
    - 11 data sets are available to the public via EDG client
    - Lower level products (L1A/B) generated in Japan and sent to EDC DAAC for archive
    - Other 9 data sets processed on-demand
  - Approximately 36 TB in archive
  - Up to 142 GB/day received on D3 tape from Japan (116 GB/day in 6/01)

Current Archive

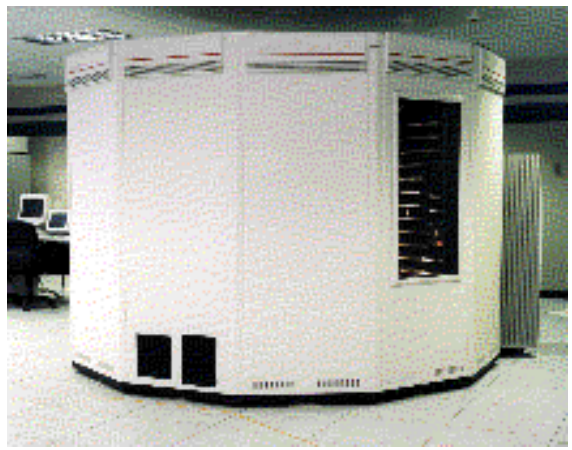
- MODIS
  - First data at EDC in February 2000
  - 25 data sets have been released to date
  - Approximately 50 TB in archive
  - Up to 308 GB/day increasing to 924 GB/day in FY03 (118 GB/day in 5/01).
    - Plan is to reprocess all the MODIS data each year.
    - EDC DAAC archives and distributes data processed at Goddard.
    - All MODIS data received via network from Goddard (EBNET)
  - AQUA (PM-1) satellite expected early CY02.
    - Mounts another MODIS instrument – part of the reason data flows increase over ’02 and ’03.

DAAC Systems

- Dedicated Version 0 systems
  - 5 SGI servers, 1 NT server
  - 2.5TB Odetics juke box
  - Stand-alone SIR-C suite of systems (moving to the archive)
  - Over 1 TB disk
- Use of shared EDC systems
  - Science data purchase

DAAC Systems

- EOSDIS Core System (ECS)
  - Five SGI Power Challenge XL
  - Three SGI Origin 2000
  - Approx. 5 TB RAID
  - Three STK Powderhorns (5,000 9940 tapes for up to 600 TB)
    - Currently migration from D3 to new media is in work.
  - Numerous SGI, HP, and SUN workstations
  - Variety of media devices



Current Ingest/Distribution Specification

- Ingest to ECS includes Landsat 7, ASTER, and MODIS (Terra and Aqua)
  - Landsat 7 154 GB/day
  - ASTER 142 GB/day
  - MODIS 154 GB/day Terra *
  - MODIS 77 GB/day Aqua *
  - MODIS 77 GB/day Terra reprocessing *
- Distribution spec is 1X ingest for most data
  - Landsat 7 55 GB/day (110 scenes)
  - ASTER 142 GB/day *
  - MODIS 308 GB/day (both Terra and Aqua) *

*These numbers scheduled to increase.

DAAC Project Management

- Staff
  - 2+ Government – Program Manager, Deputy PM, some science support
  - Approximately 100 contract staff
    - NASA-provided staff (EOSDIS Core System contract)
    - USGS-provided staff (EDC Technical Services local contract)
      - Branch-supplied staff (provided by local line organization)
      - Matrixed staff (provided by local line organizations outside the branch)
      - Service support (provided by govt. managers supported by contractors)
  - Additional NASA external support and/or collaboration
    - ECS support
    - Network support
    - EOS Data Gateway Development and support
    - Stennis, JPL, GSFC, various USGS sites, various Universities

DAAC Project Management

- Engineering and development activity
  - Concepts, estimates, proposals, and plans as required
  - Process development, ops concepts a specialty
    - From baseline process and procedures to “real” process/procedures
    - Numerous work-around and local solutions
  - Development, implementation, integration, installation, and test (or a subset thereof) of software and hardware systems
  - Flexibility to benefit customers and USGS
    - Frequently work across contracts, projects, organizations, agencies
    - USGS / EDC systems integrated with NASA systems
    - NASA systems integrated with USGS / EDC systems
    - Can and do draw on resources “outside” the DAAC

DAAC Project Management

- Operations activity
  - Staffing profile
    - From baseline to “real” profile with continued modification based on system changes
    - Flexible allocation of work to positions based on experience and tasking.
  - Training
  - Operations readiness
    - Numerous tests, activities, and reviews in preparation of operations
  - Work-arounds
    - Local solutions to system problems – short- and long-term
  - Facility
    - HVAC, space and movement of equipment, office accommodation
  - System monitoring
    - Statistical information, error reporting, etc.

DAAC Project Management

- Science and customer support
  - Customer services model
  - Tools and application support
    - Development of tool for new MODIS data via local university
    - Knowledge of data within science and user services segments
  - Science and professional organization/conferences
  - QA support with multiple science teams
  - Support validation of data and field campaigns
  - User interface support